Open File No. 124

EGSP
OPEN FILE # UGR 124

REPORT ON PHASE III OF THE MULTIVARIATE STATISTICAL CHARACTERIZATION OF GEOCHEMICAL DATA GENERATED BY THE EGSP FOR CORES 20403, 11940, 12041, GG5 AND 7239.

by

Steven M. Warshauer
Consulting Geologist

Introduction

This phase of the multivariate statistical characterization of geochemical data generated by the EGSP for cores 20403, 11940, 12041. GG5 and 7239 marks the beginning of the sample by sample, or Q-mode, analysis. This is an important phase as the positive results obtained from this phase, when coupled with those from the remaining portion, will lead to geologically significant interpretations of the abundant data at hand.

Methods

Q-mode cluster analysis was applied to the data sets obtained from phases I and II of this project. Clustering was initiated using the unweighted pair-group method with arithmetic averages as that algorithm imparts the least structural distortion (distortion is inherent anytime one displays a multidimensional structure in 2 dimensions) to the resulting dendrograms. Distortion was evaluated by the calculation of Cophenetic correlation coefficients for each of the dendrograms.

Analyses were obtained for each of the 5 cores separately as well as for 3 different combined sets. The combined data sets consisted of:

1. A 309 sample by 24 variable geochemical data set for all 5 cores. Data was expressed as weight percent of the total.
2. A 268 sample by 9 variable mineralogical data set of all the cores minus GG5 as it did not have a complete mineralogical analysis. Mineralogical data was expressed as a percentage of the total peak intensity as this closely approximates weight percent (Renton, 1978).
3. A 268 sample by 33 variable data set consisting of the standardized values for the geochemical and mineralogical data for all of the cores minus GG5.

Cophenetic correlation values were obtained from the resulting 9 dendrograms. Table 1 is a listing of these values and demonstrates, due to the fact that all of the values approach 1.0, the low level of distortion obtained throughout the analysis. Characterization of the clusters was then accomplished by the calculation of mean variable abundances for each of the clusters. Constancy and fidelity values were abandoned as their binary based nature rendered them useless for this type of ordinal data; i.e., the important variables occurred in virtually all of the samples, differing only in abundances.

Results

Examination of the resulting 9 dendrograms demonstrated several important factors; firstly, the 5 separate dendrograms tend to have many outliers, that is, a large portion of the samples fail to cluster. Geologic interpretation would therefore be incomplete as many of the samples would not be included in the discussion.

Table 1

Cophenetic correlation values for the eight Q-mode dendrograms

Dendrogram	Cophenetic Correlation Value
1. Core 11940	0.972
2. Core 12041	0.964
3. Core GG5	0.813
4. Core 7239	0.963
5. Core 20403	0.792
6. Total data, Geochemistry	0.932
7. Total data minus GG5, Mineralogy	0.930
8. Total data minus GG5 Mineralogy + Geochemistry	0.949

More importantly, each core would have to be treated as a separate entity and an overall integration of the data would be almost impossible. Two of the combined core data sets, the geochemical and mineralogical, did have a high degree of samples that clustered (over 90%) and provide a much more geologically meaningful picture as all of the samples were compared to one another. Due to these facts the five separate dendrograms and the combined standardized geochemical plus mineralogical data sets were abandoned as uninterpretable and further discussion will deal solely with the 2 remaining combined core data sets.

Examination of the 309 sample by 24 variable total geochemistry data dendrogram indicates the existence of 7 major clusters, 4 subclusters, and a small number of samples that fail to cluster. In addition, the clusters are related at a very high level of similarity (low dissimilarity), indicating an overall sameness to the samples. Table 2 is a listing of the clusters and the variables that characterize each of them. The 268 sample by 9 variable mineralogical data set can be visually divided into 6 clusters, 3 subclusters, and a small number of outliers. Again note the overall similarity of the samples to one another. A list of the clusters and the variables that characterize each of them can be found in Table 3.

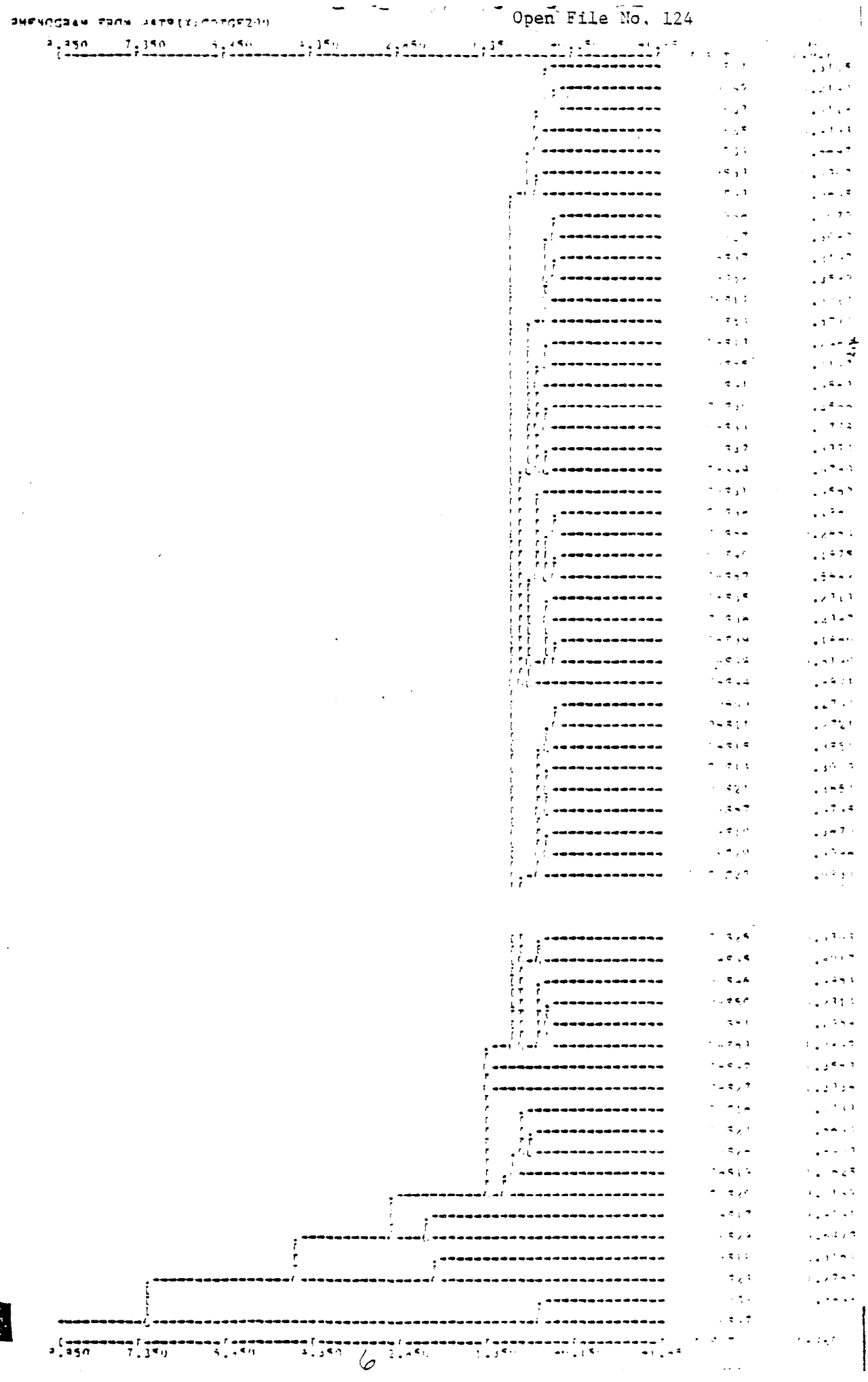
Conclusions

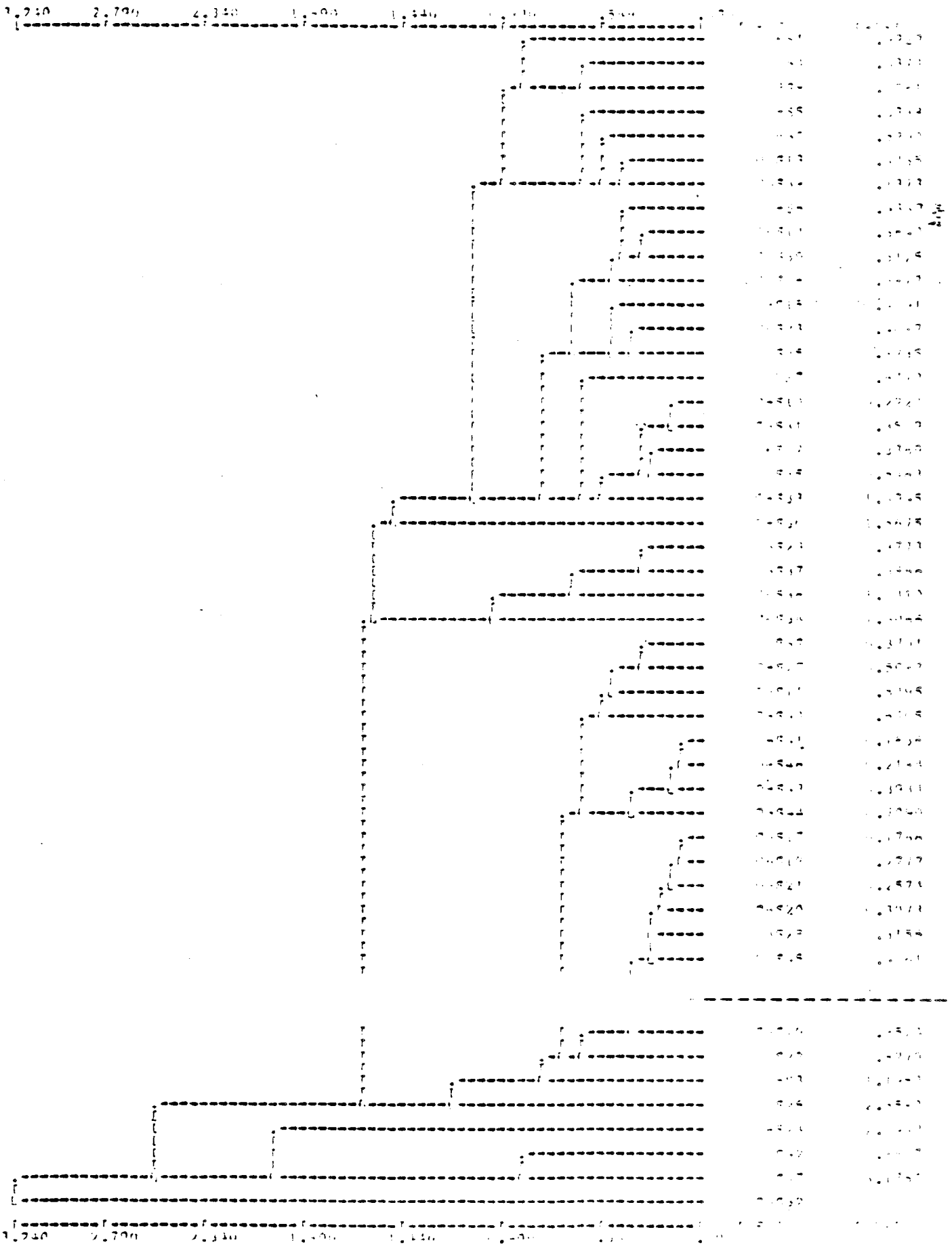
Due to the high level of similarity (low dissimilarity) among the samples it can be concluded that there are no strong differences among the shale samples dealt with in this study. It must be noted, however, that strong differences may exist between variables that were not measured (such as Uranium).

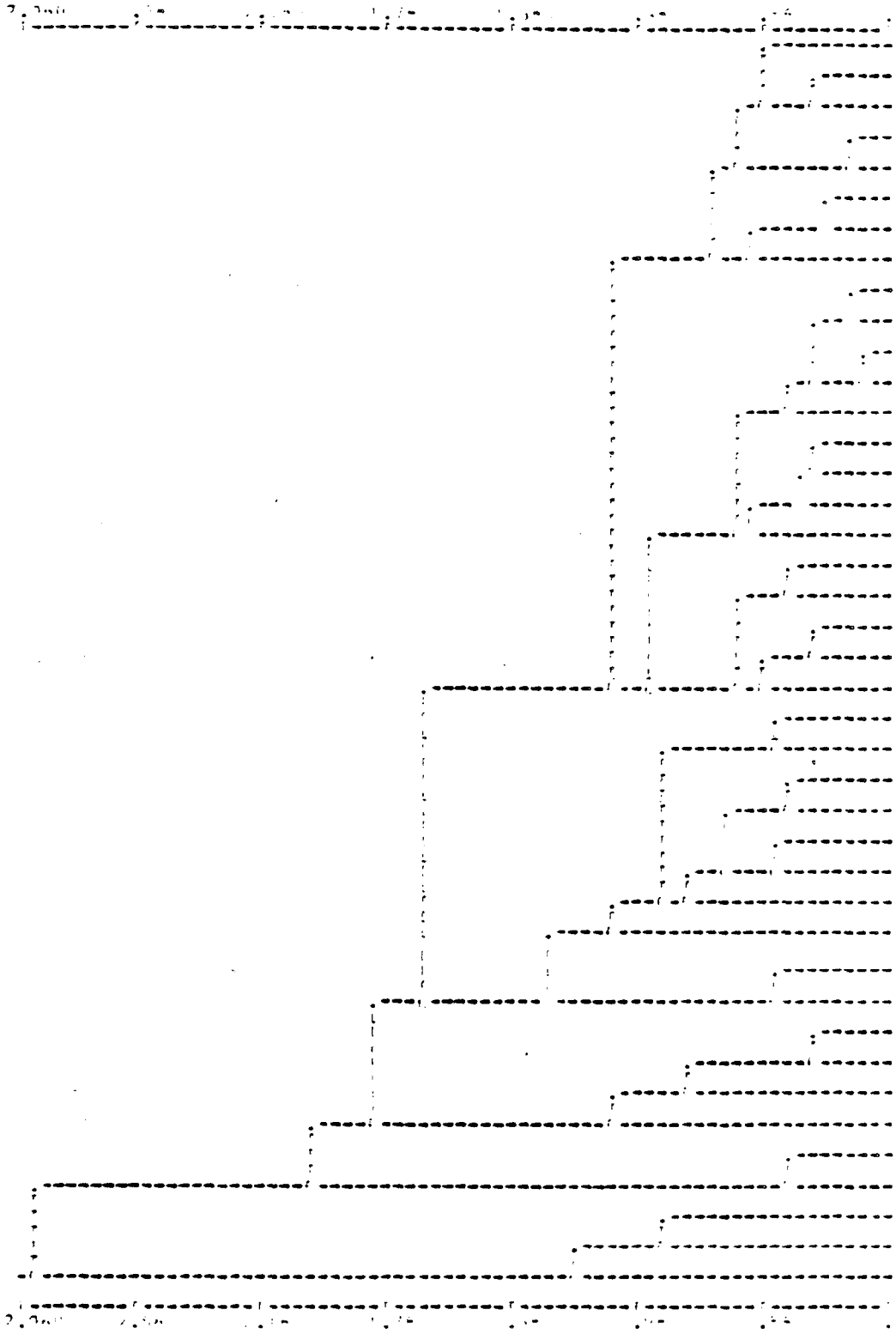
We can separate the samples into groupings based on small differences in their variable compositions.

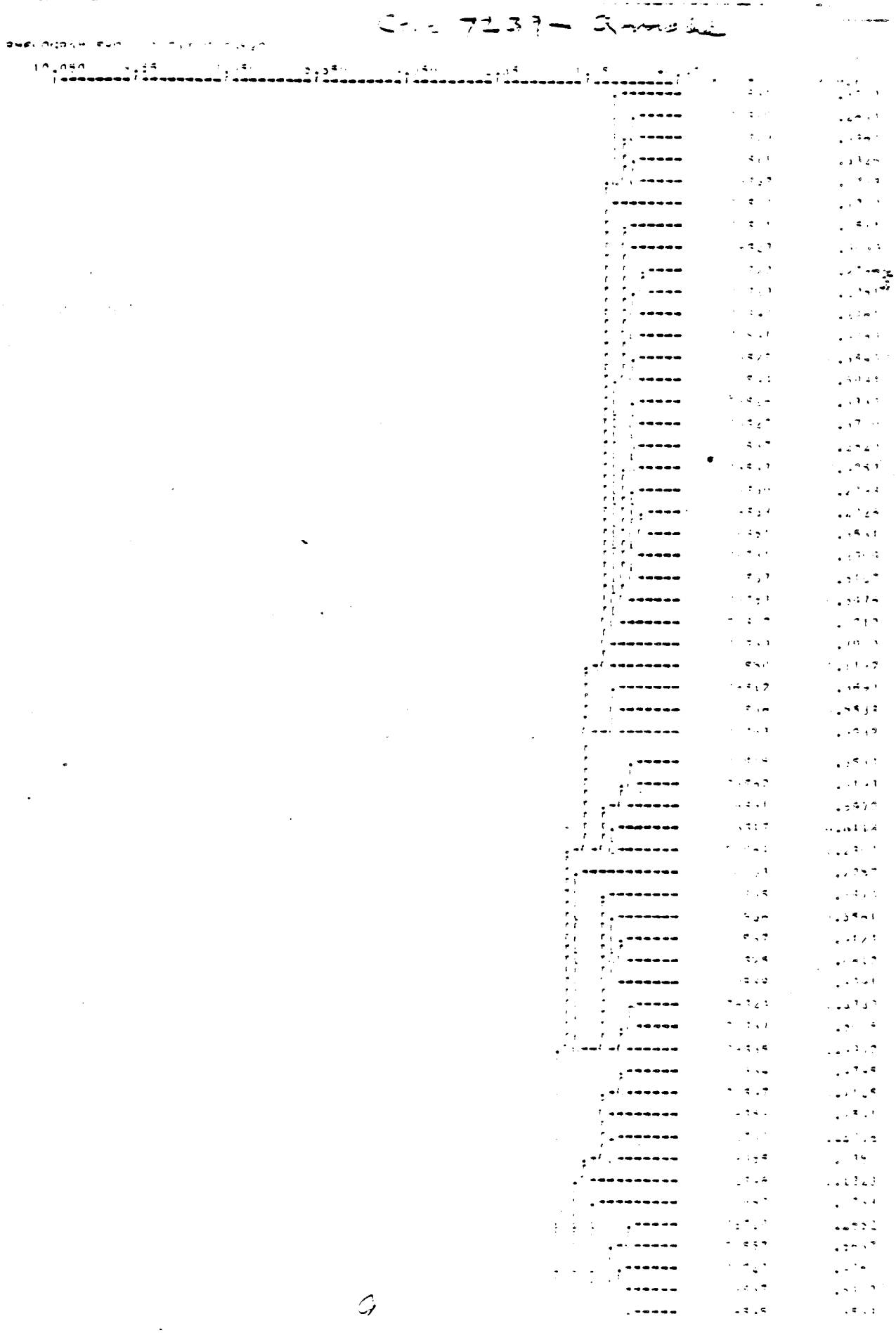
Some vertical and lateral segregation of the samples is apparent from visual inspection of the dendrograms. Whether or not there are any systematic trends to the vertical and lateral variations will become apparent in Phase IV, when the clusters will be plotted on the stratigraphic sections.

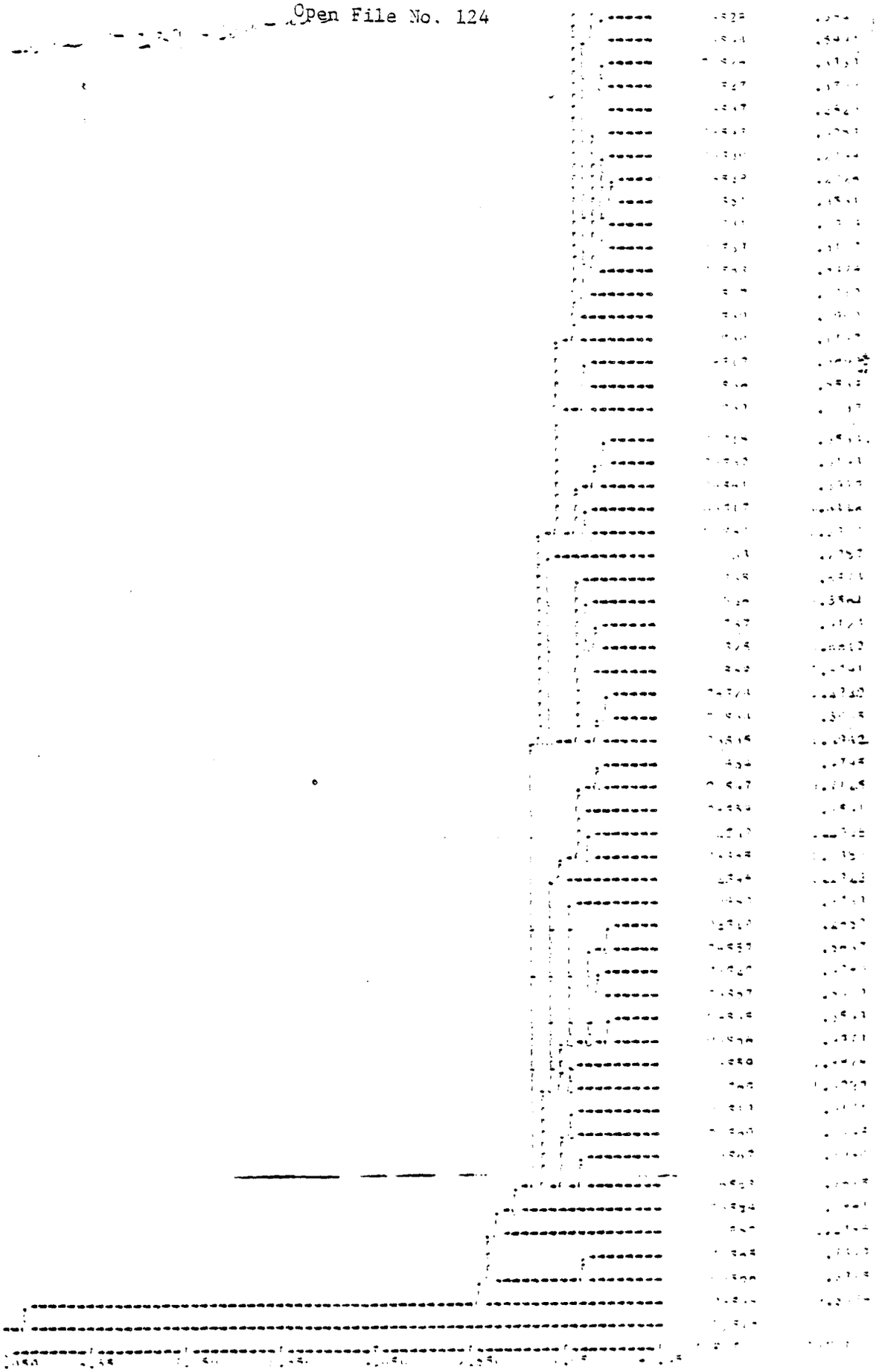
Steven M. Warshauer

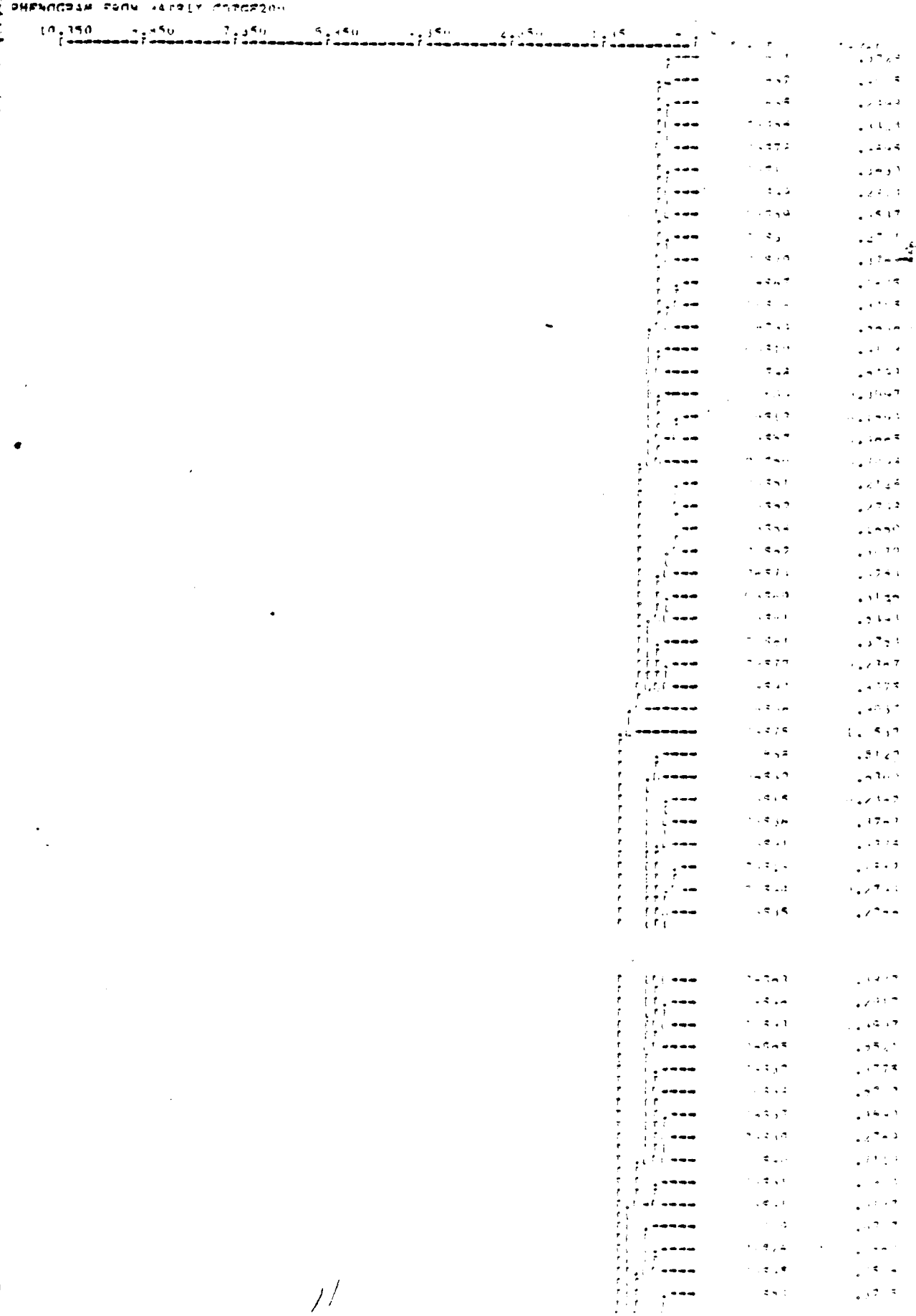


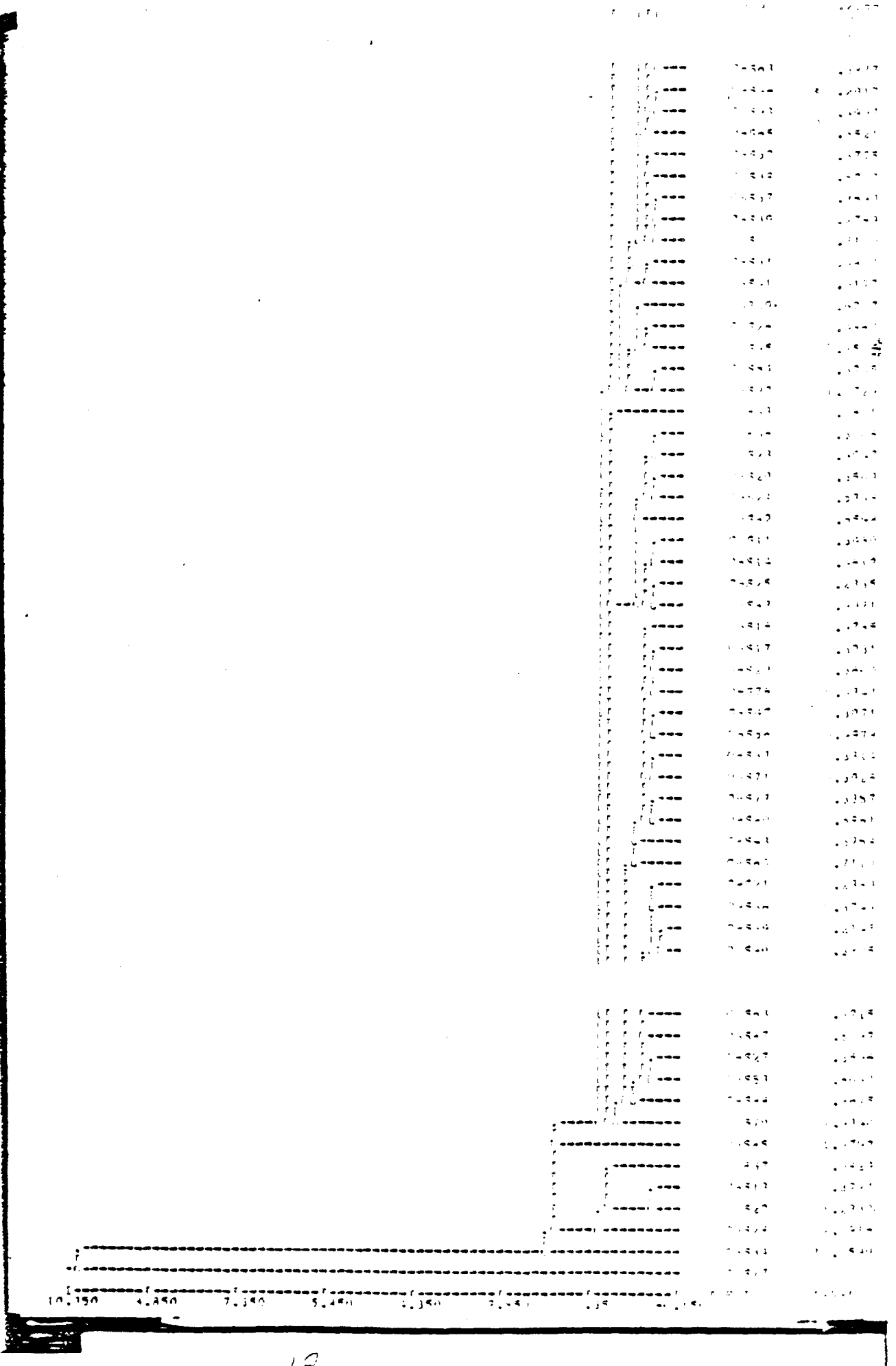

PHENOGRAM FROM MATRIX CATCH
3.240
2.790
2.340
1.890
1.440
.540



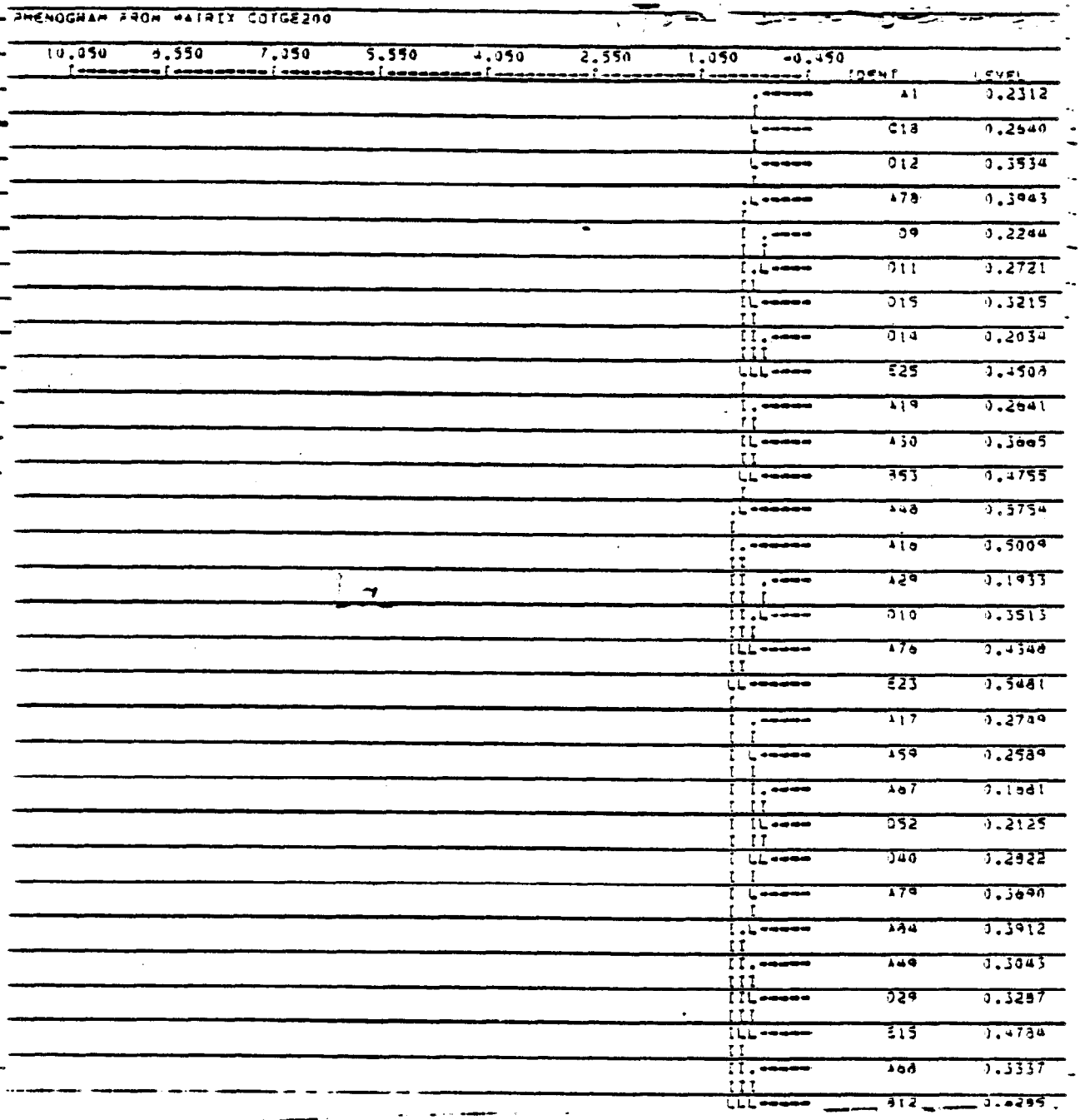





PHENOGRAM FROM MATRIX
10.350
7.350
5.450



DRENOGRAM FROM MATRIX COTGE200



	14	0.4232
	410	0.3564
	442	0.7792
	121	0.2974
	186	0.2512
	512	0.3011
	189	0.2207
	190	0.3870
	180	0.3540
	291	0.1986
	539	0.2586
	022	0.4384
	164	0.4863
	153	0.3423
	171	0.4753
	516	0.5214
	127	0.3929
	01	0.4430
	147	0.3112
	173	0.2831
	057	0.3518
	021	0.3933
	262	0.5085
	195	0.5430
	132	0.3820
	021	0.2900
	544	0.4054
	541	0.1836
	546	0.2194
	543	0.5024
	154	0.3423
	192	0.5698

PHENOGRAM FROM MATRIX COTGE200

II .-----	A38	0.2432
II I		
II.L-----	E40	0.4366
III		
III.-----	A53	0.3735
IIII		
IIIL-----	B15	0.2777
IIII		
LLLL-----	C20	0.7386
I		
I .-----	A56	0.2377
I I		
I L-----	D53	0.3116
I I		
I .L-----	C2	0.3754
I I		
I I.-----	D50	0.2314
I II		
I LL-----	D51	0.5019
I I		
I L-----	A88	0.3944
I I		
I I.-----	B63	0.3382
I II		
I.LL-----	E6	0.5821
II		
II.-----	C27	0.3943
III		
ILL-----	C31	0.6566
II		
II .-----	A87	0.2894
II I		
II L-----	C17	0.2662
II I		
.LL-L-----	C19	0.3681
I		
I .-----	A9	0.5711
I I		
I I.-----	A26	0.4119
I II		
I.LL-----	D26	0.6866
II		
II .-----	A43	0.2867
II I		
II .L-----	D45	0.4733
II I		
II.L-----	D25	0.5706
III		
LLL-----	A60	0.4212
I		
.L-----	A70	1.0049
I		
I .-----	A2	0.2549
I I		
I L-----	D5	0.3136
I I		
I .L-----	D1	0.4018
I I		
I I.-----	C14	0.2655
I II		
I II.----	D2	0.2190
I III		
I LLL----	D3	0.4752
I I		
I .L-----	D4	0.5883

	A61	0.3403
	C22	0.3944
	A72	0.2039
	C15	0.3053
	A94	0.6781
	A5	0.1980
	D54	0.2815
	A55	0.3239
	B31	0.2702
	B37	0.2560
	D37	0.3790
	A51	0.1965
	D58	0.2342
	A81	0.3241
	D6	0.2070
	D7	0.3593
	A52	0.1943
	B43	0.2686
	A74	0.3043
	A58	0.1901
	A62	0.4523
	A69	0.3599
	B21	0.4852
	B26	0.3339
moderate hydrogen	B27	0.5820
	B22	0.2086
moderate-Hy	B29	0.2951
	B40	0.2461
2 [illegible]	B41	0.4116
	B28	0.3540
heavy [illegible] sulfur	B44	0.4043
	D43	0.4517

I IIL------	823	0.4954
I III		
I III.-----	049	0.3572
I IIII		
I ILLL-----	637	0.6164
I II		
I LL-------	033	0.7359
I I		
I L--------	475	0.7509
I I		
I L--------	939	0.3442
II .----	412	0.1763
II I		
II .L----	457	0.3096
II I		
II I.----	038	0.1860
II II		
II LL----	056	0.2602
II I		
II L-----	635	0.3575
II I		
II .L-----	933	0.4617
II I		
II I.-----	450	0.2676
II II		
II IL-----	041	0.3706
II II		
II IL-----	830	0.2498
II II		
II II.----	438	0.2026
II III		
II ILL----	351	0.3491
II II		
II LL-----	030	0.4278
II I		
II I.-----	832	0.2582
II II		
II IL-----	013	0.2462
II II		
II LL-----	055	0.4032
II I		
II .L------	034	0.5544
II I		
II I .-----	824	0.3339
II I I		
II I L-----	834	0.3617
II I I		
II I L-----	035	0.2313
II I I		
II I L-----	036	0.2382
II I I		
II I I.----	039	0.1860
II I II		
II I.LL----	048	0.4334
II II		
II IL------	633	0.4844
II II		
II LL------	335	0.6119
II I		
II L-------	31	0.5614
II I		
II I .-----	310	0.2644
II I I		
II I.L-----	319	0.3810

D44	0.4408
B11	0.3596
D32	0.2766
E36	0.7259
C34	0.8397
B50	0.9111
A66	0.6224
B14	0.6913
C32	1.1389
C36	1.2271
A8	0.5188
A37	0.2654
E28	0.4115
A42	0.3402
E17	0.1766
E18	0.2272
E21	0.2574
E20	0.3996
D16	0.2406
E45	0.4884
C11	0.3440
E19	0.5961
A15	0.2536
A36	0.3271
A91	0.3781
A18	0.1890
A44	0.2933
A35	0.2369
A63	0.4129
A46	0.2813
A93	0.3632
C5	0.2872

3

moderately high

Iron

C9	0.2011
E27	0.3432
E2	0.5023
B17	0.3224
E11	0.5275
A39	0.2958
A40	0.3947
C39	0.4593
C7	0.3049
C12	0.2301
E22	0.5527
B64	0.5941
B67	0.3446
C40	0.5668
C10	0.3804
C13	0.4238
C38	0.7692
A31	0.4322
A41	0.8709
E3	1.0578
A45	0.4855
D24	0.6636
B69	0.8051
B13	0.6195
D19	0.9777
B60	1.1407
D20	0.5419
E26	1.5356
A3	0.6629
C3	0.4349
D27	0.8960
A6	0.2752
E10	0.2224

4 Low Pyr. Sul. High Fe O

E31	0.3611
A23	0.4336
D31	0.2973
E12	0.5320
B36	0.3832
E7	0.4312
C35	0.5275
D46	0.5831
A82	0.3922
C16	0.4522
C26	0.6617
A11	0.3536
A14	0.3150
E8	0.5375
A20	0.3255
A24	0.4192
A25	0.2346
E13	0.2739
A83	0.6193
E38	0.7328
C6	1.0200
A28	0.7777
E5	0.3433
E9	1.0709
A7	0.6690
E4	0.9785
A13	0.3404
A22	0.4344
C4	0.3000
E1	0.5735
C3	1.2355
B5	0.6345

5 Highest
Ex. Ex. A1

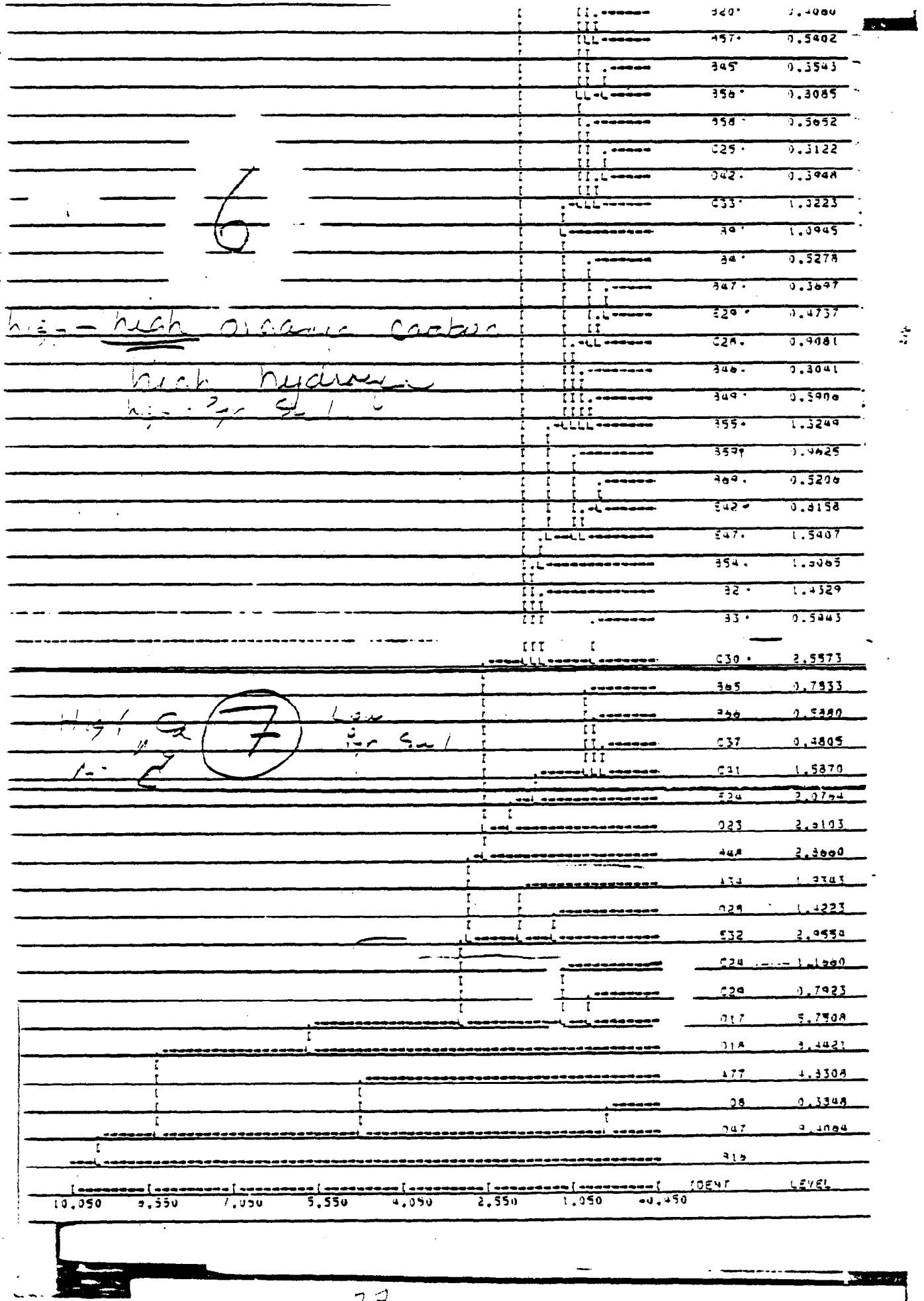
C23	0.7676
E30	0.8247
B6	0.5561
B7	0.4124
B25	0.6612
B8	1.7473
A85	0.5709
E34	0.7130
B18	0.2652
B52	0.5637
B20	0.4060
B57	0.5402
B45	0.3543
B56	0.3085
B58	0.5652
C25	0.3122
D42	0.3948
C33	1.0223
B9	1.0945
B4	0.5278
B47	0.3697
E29	0.4737
C28	0.9081
B46	0.8041
B49	0.5906
B55	1.3249
B59	0.9625
B69	0.5206
E42	0.8158
E47	1.5407
B54	1.6065
B2	1.4329
B3	0.5443
C30	2.5573
B65	0.7833
B66	0.5880
C37	0.4805
C31	1.5870
E24	2.0764
D23	2.6103
B48	2.3660
A34	1.9343
D28	1.4223
E32	2.9554
C24	1.1660

6

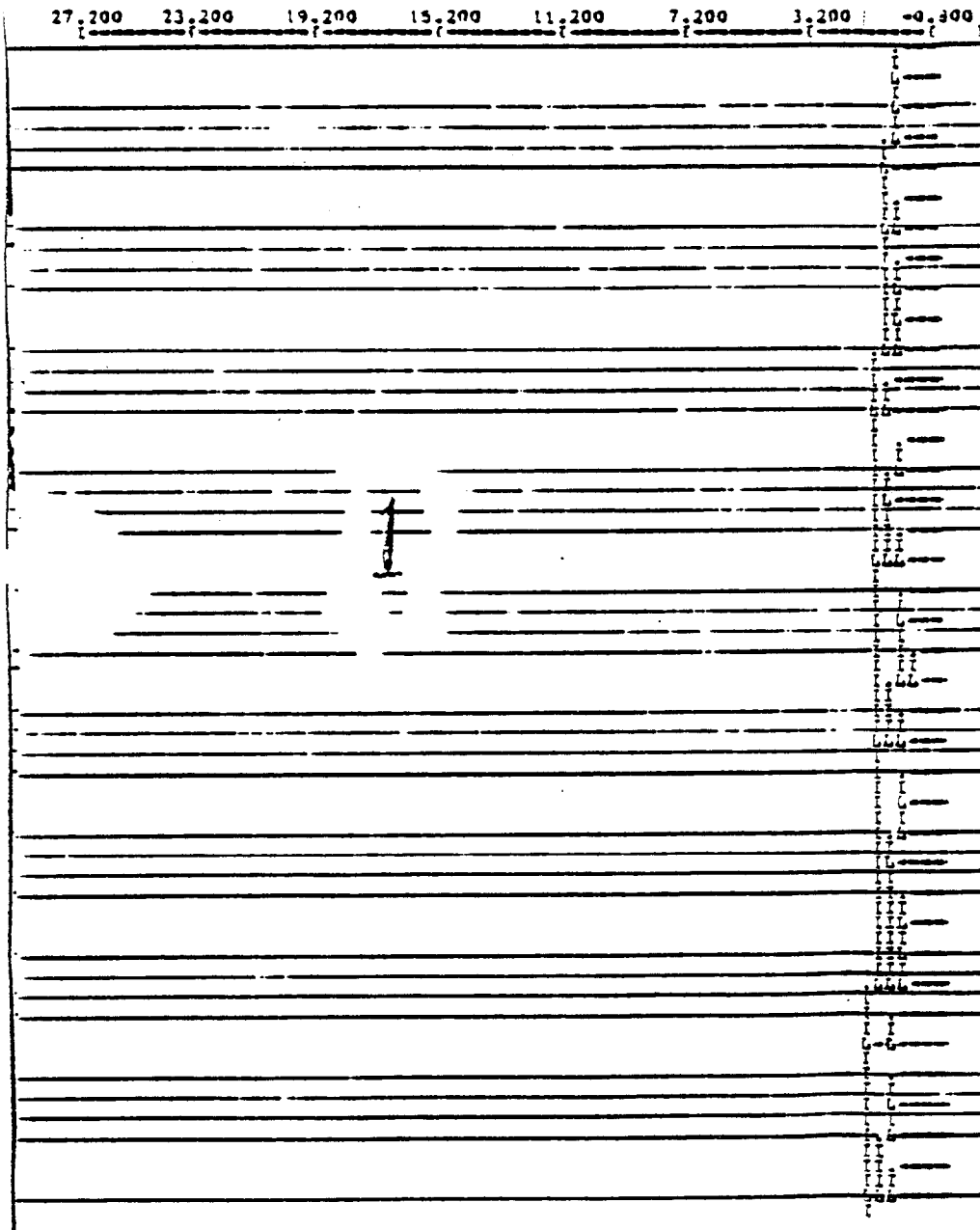
high organic carbon

high hydrogen

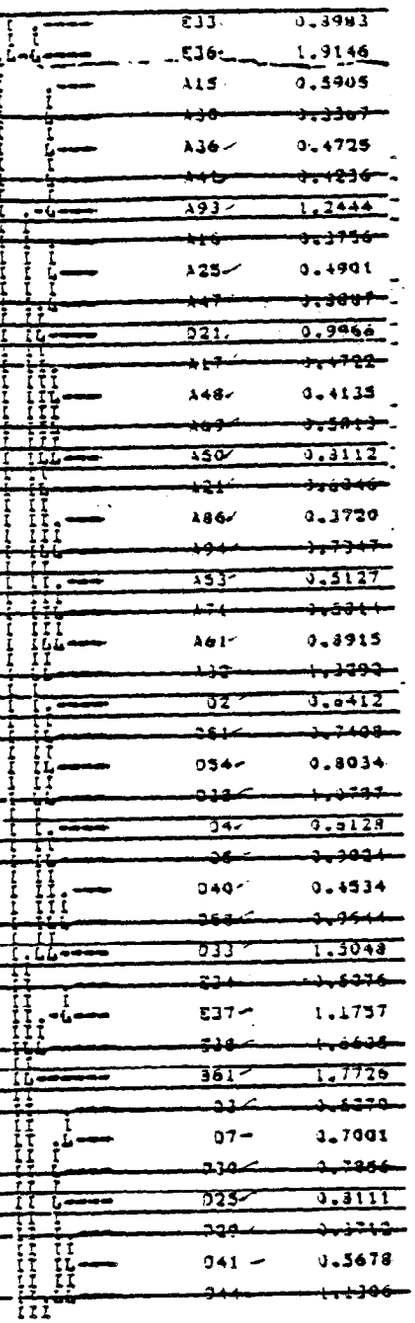
7


6
high organic carbon
high hydrogen
7
IDENT
LEVEL
10.050
9.550
7.050
5.550
4.050
2.550
1.050
-0.450

PHENOGRAM FROM MATRIX COTGE200



IDENT	LEVEL
A1	0.4387
A62	0.3004
A81	0.5343
A74	0.6427
A39	0.3359
A33	0.4955
D24	0.3537
A58	0.4494
A67	0.5031
A68	0.5888
A79	1.0880
A19	0.8731
B1	1.0992
A13	0.5606
A84	0.6370
D35	0.7079
A51	0.4468
D32	1.2878
A52	0.4479
A77	0.2693
A79	0.1948
A85	0.7089
A76	0.5698
D31	1.0403
D22	0.5944
D42	0.4989
D56	0.7157
D53	0.9200
D37	0.2539
D48	0.3678
D39	0.4860
D55	1.4857
A54	0.7989
A80	1.5488
D45	0.6368
D50	0.9258
D46	1.3367
E29	0.6509
E31	1.4342



E33	0.3983
E36	1.9146
A15	0.5905
A38	0.3367
A36	0.4725
A41	0.4236
A93	1.2444
A16	0.3756
A25	0.4901
A47	0.3887
D21	0.9966
A17	0.4722
A48	0.4135
A69	0.5813
A50	0.3112
A21	0.6346
A86	0.3720
A94	0.7347
A53	0.5127
A74	0.5814
A61	0.8915
A32	1.3293
D2	0.6412
D61	0.7408
D54	0.8034
D38	1.0787
D4	0.5128
D6	0.9024
D40	0.4534
D58	0.9644
D33	1.5048
E34	0.6376
E37	1.1757
E38	1.6635
B61	1.7726
D3	0.6270
D7	0.7001
D30	0.7856
D25	0.8111
D29	0.3712
D41	0.5678
D44	1.1306

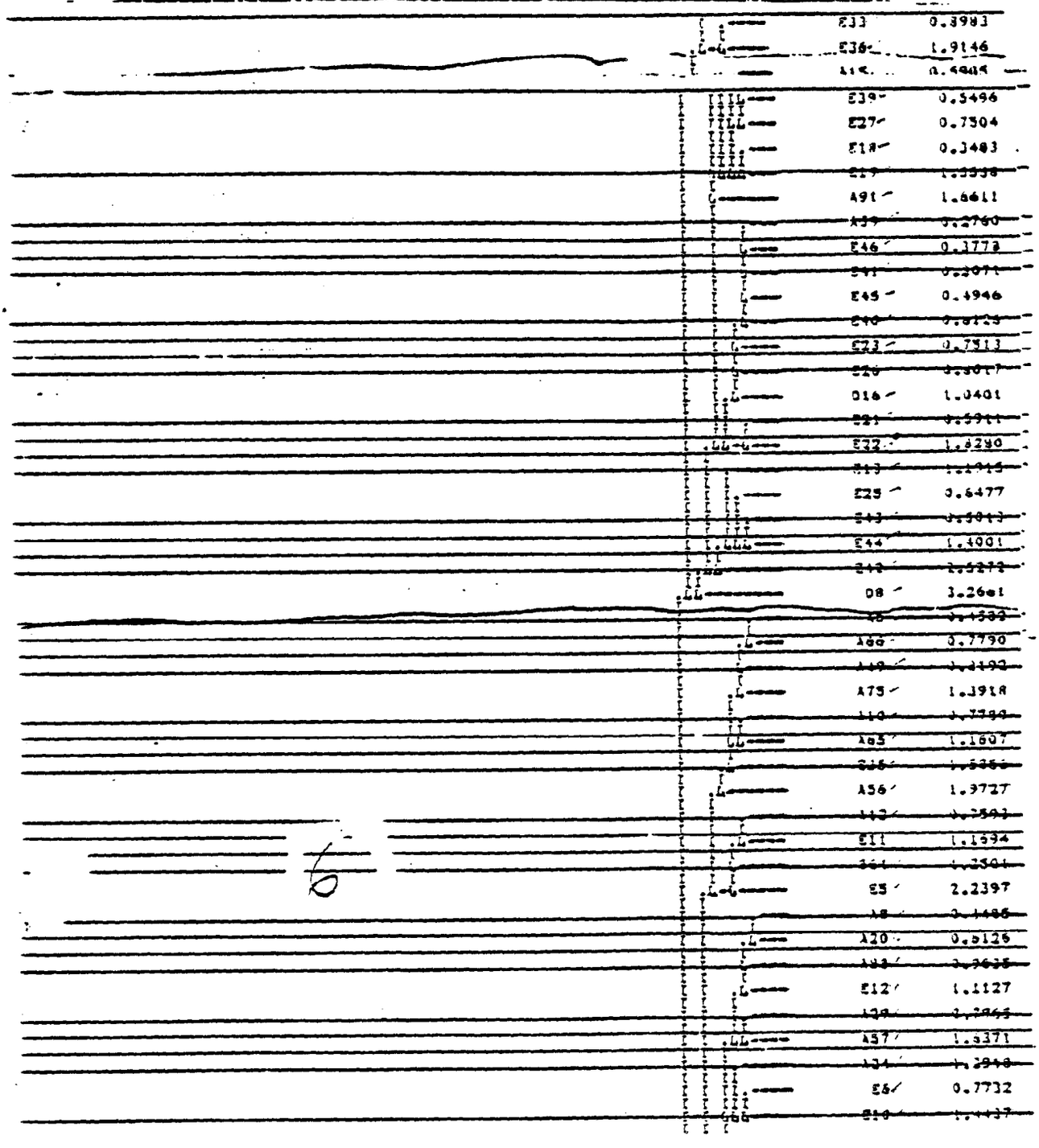
2

012	0.6749
049	0.7391
019	1.3432
09	0.7993
013	1.0579
215	1.9838
A26	0.3469
A43	0.9225
A92	1.4649
026	2.3293
AM3	0.5014
95	0.7819
B3	0.5774
959	0.9367
B10	0.5774
941	1.2987
B4	1.1840
948	0.3253
B56	1.3840
96	0.5735
B7	0.7437
925	0.3810
B47	0.9825
98	2.6000
A27	1.5096
A31	0.4643
A38	0.7558
A40	0.3537
A90	0.3405
487	0.6031
A89	1.2599
911	0.6489
D14	0.9110
945	1.3036
E2	0.7036
27	0.7982
E4	0.2853
E16	0.5976
E20	0.6684
917	0.9743
E14	0.4664
928	0.3532

3

4

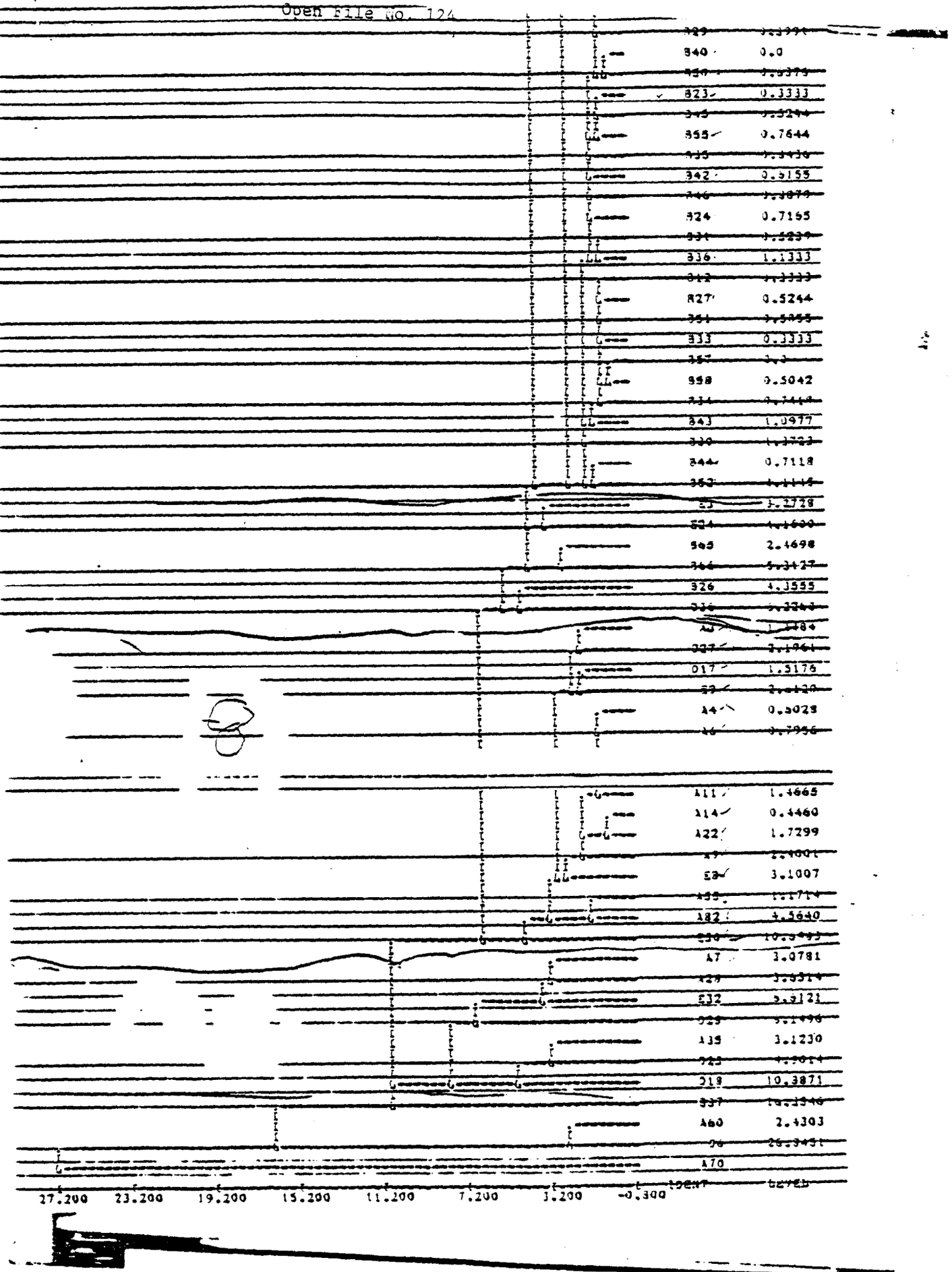
5


E33 0.3983
E36 1.9146
E39 0.5496
E27 0.7504
E18 0.3483
A91 1.6611
E46 0.3778
E45 0.4946
E23 0.7513
D16 1.0401
E25 0.6477
E44 1.4001
D8 3.2661
A60 0.7790
A75 1.3918
A85 1.1607
A56 1.9727
E11 1.1594
E5 2.2397
A20 0.5126
E12 1.1127
A57 1.6371
E6 0.7732

6

052	2.0637
034	0.5981
047	1.0946
043	2.3600
E1	3.6048
A2	0.4478
A45	0.5123
A46	0.5877
A72	0.8050
B60	0.8570
A73	0.5581
B11	0.4041
B54	0.4470
B49	1.1064
B62	0.7922
B63	0.5774
B69	1.2513
B9	0.5591
B16	0.4203
B19	0.6035
B21	0.4714
B53	1.5581
A23	0.5375
A42	0.7032
A37	0.8953
A34	0.5552
A44	1.1598
B1	0.6438
D57	1.2348
B67	0.9843
D10	2.0552
B20	1.3507
E47	2.1512
A18	1.0258
B69	1.3441
A63	0.4955
B17	0.5859
B20	0.7962
A64	0.5090
B14	0.8194
B15	0.3859
B18	0.6217
E31	1.4342

E33	0.3983
E36	1.9146
A15	0.5905
839	1.680
813	1.1514
822	0.6646
828	0.5011
832	0.7563
830	1.9315
32	0.2867
829	0.3991
840	0.0
850	0.6375
823	0.3333
845	0.5244
855	0.7644
835	0.8436
842	0.6155
846	1.8870
824	0.7165
831	0.5239
836	1.1333
812	0.3333
827	0.5244
851	0.5955
833	0.3333
857	0.0
858	0.5042
834	0.7419
843	1.0977
830	1.3723
844	0.7118
852	4.1145
63	3.2729
824	4.1600
865	2.4698
866	5.3127
326	4.3555
336	4.2263
A3	1.1484
327	2.1961
017	1.5176
69	2.6120
A4	0.6028
A6	0.7956
A11	1.4665
A14	0.3460


840 0.0
823 0.3333
855 0.7644
342 0.5155
824 0.7165
836 1.1333
827 0.5244
833 0.3333
858 0.5042
843 1.0977
844 0.7118
2.2728
565 2.4698
326 4.3555
017 1.5176
A4 0.5028
A11 1.4665
A14 0.4460
A22 1.7299
53 3.1007
A82 4.5640
A7 3.0781
532 3.9121
A35 3.1230
519 10.3871
A60 2.4303
A70
27.200
23.200
19.200
15.200
11.200
7.200
3.200
-0.300

3½

Table 2

Geochemical variables that best characterize the clusters generated by the Q-mode analysis of the total data set

Cluster	Characterizing Variable(s)
1	none (standard cluster)
2	Pyritic sulfur
3	Iron
4	Magnesium
5	Silica
6	Organic Carbon
7	Calcium-Magnesium

Table 3

Mineralogical variables that best characterize the clusters generated by the Q-mode analysis of the total data set minus Core GG5.

Cluster	Characterizing Variable(s)
1	Pyrite
2	Orthoclase
3	Chlorite-Siderite
4	no Siderite
5	Illite
6	Quartz-Plagioclase-Calcite-Dolomite